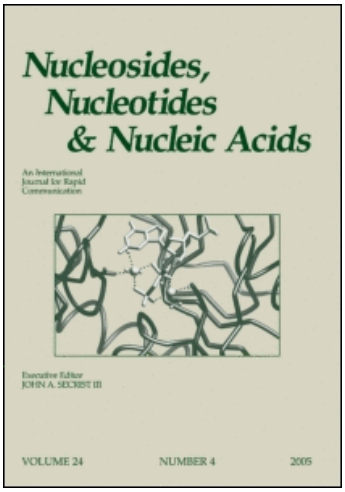

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Free Energy Calculations Predict Sequence Specificity In DNA-drug Complexes

T. Härd[a]; L. Nilsson[a]
[a] Department of Medical Biophysics, Karolinska Institutet, Stockholm, Sweden

To cite this Article Härd, T. and Nilsson, L.(1991) 'Free Energy Calculations Predict Sequence Specificity In DNA-drug Complexes', Nucleosides, Nucleotides and Nucleic Acids, 10: 1, 701 — 709
To link to this Article: DOI: 10.1080/07328319108046577
URL: http://dx.doi.org/10.1080/07328319108046577

NUCLEOSIDES & NUCLEOTIDES, 10(1-3), 701-709 (1991)

FREE ENERGY CALCULATIONS PREDICT SEQUENCE SPECIFICITY IN DNA-DRUG COMPLEXES

T. Härd and **L. Nilsson**, Department of Medical Biophysics, Karolinska Institutet, S-10401 Stockholm, Sweden.

Abstract. Molecular dynamics simulations of DNA-netropsin complexes in water were performed using the thermodynamic cycle-perturbation method to calculate the free energy difference between complexes with an adenine-containing binding site and corresponding complexes where adenines are replaced by 2,6-diaminopurines (dap). The calculations predict a free energy difference of 3.7±0.9 kcal/mol (at 300K) in favour of netropsin binding to an $(AATT)_2$ DNA sequence compared to a $(dapdapTT)_2$ sequence.

Introduction

The possibility to theoretically predict the effects of small structural changes on sequence-specific DNA-ligand equilibria would be very valuable in understanding the mechanisms of gene regulation or for designing sequence specific DNA-binding drugs. Recently, free energy calculations based on the thermodynamic cycle-perturbation approach and molecular dynamics (MD) simulations have been successfully employed to theoretically 'reproduce' experimentally observed differences in binding constants in various host-guest systems including enyme-substrate complexes[1], suggesting that similar calculations might be applied also to DNA complexes. The DNA-netropsin complex provides an excellent model system to test this possibility for two reasons: it is very well characterized with regard to structure[2] and thermodynamics[3] and there are large (free energy) differences between complexes involving different DNA sequences.

We have performed calculations of the free energy difference between two DNA-netropsin complexes (with netropsin bound to an AATT site) and the corresponding model of the active site of the *Tetrahymena* intron based on principles of RNA folding derived from a crystallographic structure of $tRNA^{Phe}$ [11]. Michel *et al.*[2] were in a recent study able to identify the parts of the active site of the intron that recognize the guanosine cofactor which initiates the 5' cleavage reaction.

It is likely that the three-dimensional structures of the catalytic site of self-splicing introns will have several common features, especially given the high homology between secondary structural elements of group 1 and group 1A introns. Theoretical and simulation methods can give a very detailed description of these systems, not only to fill in the gaps between the experiments, but also to provide predictions and suggestions for further studies. We have therefore extended the studies of Cech and coworkers[1,12] by model-building the structural elements of the *nrdB* intron into the proposed structural model of the *Tetrahymena* active site and subsequently performed a molecular dynamics simulation with the model of the T4 *nrdB* intron[13]. A molecular dynamics simulation gives a more efficient relaxation of the model than what can be achieved by energy minimization, and also provides some insight in the stability and flexibility of the various parts of the model.

Methods

The building of the three dimensional arrangement of the helices, the placement of joining single stranded segments and the docking of the arginine and guanosine cofactors were performed using the program HYDRA (Polygen, Waltham, MA) on an IRIS 3120 Workstation (Silicon Graphics Inc., Mountain View, CA). Single stranded segments were built by sequential docking of single nucleotides, adjusting torsion angles of the nucleotides as necessary, from both ends of the stretch that was to be covered.

Molecular dynamics simulations were performed using the program CHARMM[14] with energy parameters for nucleic acids[15], as described in more detail by Nilsson *et al.*[13].

A reduced system consisting of 111 nucleotides with 111"hydrated" sodium ions Singh *et al.*[16] were included in the simulations. The ends of the (artificially) cut helices P4, P6 and P7, as well as the base pairs G_{500}–C_{36} and U_{41}–G_{495} at the ends of P3 were further kept from fraying by the addition of harmonic constraints to keep the Watson-Crick basepairing intact.

Following an initial energy minimization molecular dynamics simulations were carried out with a time step of 2 fs using Langevin dynamics for 50 ps at 500K. An average structure for the 45-50 ps

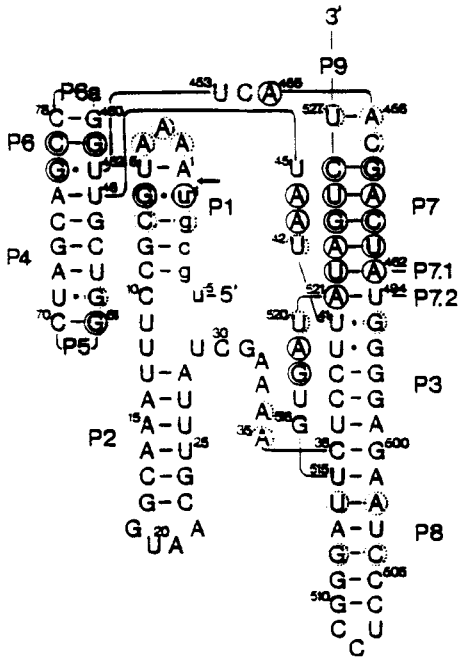

FIG.1. Secondary structure model of the *nrdB* IVS with helical elements positioned according to Shub *et al.*[17] and Kim & Cech[1]. *Filled circles* denote highly conserved nucleotides within group 1 and 1A introns, *dotted circles* denote nucleotides conserved only among the bacteriophage T4 introns.

period was energy minimized and used for the final analysis. The arginine inhibitor was docked to the average structure, and the new complex was again energy minimized.

Results and discussion

The initial three dimensional model of the T4 *nrdB* intron was based on the secondary structure model of Shub *et al.*[17]. A two dimensional arrangement of the secondary structure regions P1, P2, P3, P4, P6, P7 and P8 was constructed (FIG. 1) based on the *Tetrahymena* IVS model suggested by Kim & Cech[1]. Three dimensional helical segments

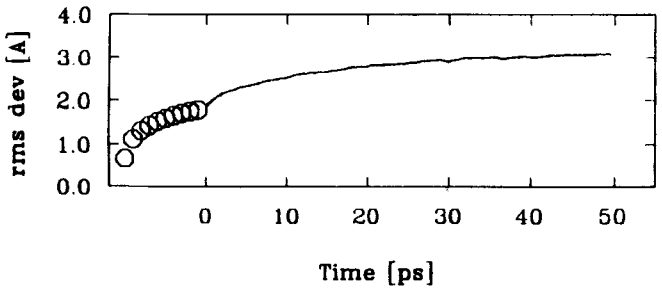

FIG. 2. Overall RMS deviation from the initial structure as a function of time in the MD simulation. The circles at negative time represent intermediate structures in the initial minimization (separated by 100 minimization cycles).

consisting of regions P1-P2, P7-P3-P8, and P4-P6 were then generated using A-RNA helix parameters[18] and arranged in three dimensions[13] as suggested by Kim & Cech[1].

The cofactor guanosine was initially docked to the splice site in P1 in two ways: First with the guanosine base close to the base pair U_{-1}-G_6, which was rearranged to allow the formation of a hydrogen bonded base triplet as suggested by Inoue *et al.*[19], and secondly in a fashion proposed by Michel *et al.*[2] with the guanosine base interacting with the basepair G_{458}-C_{526} in helix P7. Both of these structures were energy minimized and simulated for 50 ps. The first mode of binding leaves the 3' hydroxyl group of the cofactor >10Å from the actual cleavage site in P1. This simulation was therefore discarded in the final analysis presented below.

The energy minimization relieved the worst clashes that had been introduced in the modelling, thus decreasing the energy of the system, with a small change in the overall structure. During the course of the MD simulation the structure continued to relax (FIG. 2), increasing the overall RMS deviation from the initial model built structure to 3 Å. The potential and kinetic energies of the system during the MD simulation shown in FIG. 3 indicate that the system is stable energetically and structurally (FIG. 2) at the end of the run.

The majority of the nucleotides remained within 2-3Å of their starting positions, with a small increase in deviation per nucleotide

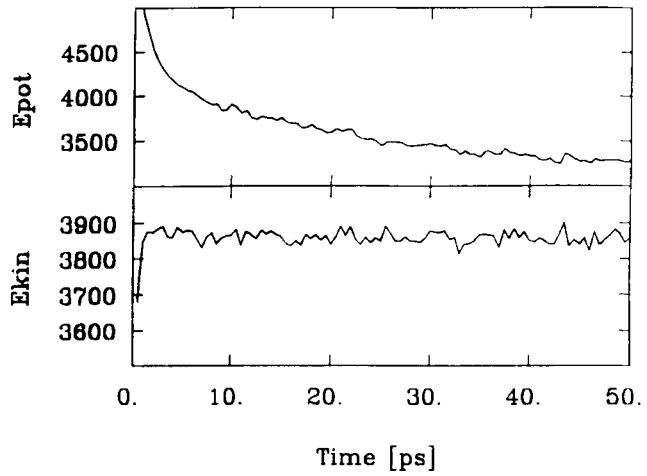

FIG. 3. Potential and kinetic energy as a function of time in the MD simulation.

towards the end of the simulation. Nucleotides C_9, A_{34}, U_{494}, and G_{496} are the ones that deviate the most (>4.5Å), whereas nucleotides U_{26}, A_{456} and C_{457} show the smallest deviations (<1.2Å)

A closer examination of the individual helices reveals that although they all retain their helical character and overall shape, there are small readjustments of nucleotide positions and base pairs within the helices, still leaving the overall helical structure intact and in place relative to the other elements in the system. One example is A_{521} which in the final structure has turned almost 90^0 so that it no longer is stacked between helices P3 and P7. This is not too surprising since this is one of the stress points in the structure, where the phosphate backbone leaves helix P7 and continues as a single strand past P3 before joining P8 again (FIG. 1).

The suggestion by Michel *et al.*[2] that the cofactor interacts with a basepair at the end of helix P7 in the group 1 *Tetrahymena* IVS was carried over to the group 1A *nrdB* IVS model. In the *nrdB* system the guanosine should form hydrogen bonds to G_{458}, which is basepaired to C_{526}, at the corresponding location in helix P7. This brought the O3' of the cofactor within 5Å of the site of the initial cleavage reaction in helix P1 at the phosphodiester bond between A_1 and U_{-1}, without any

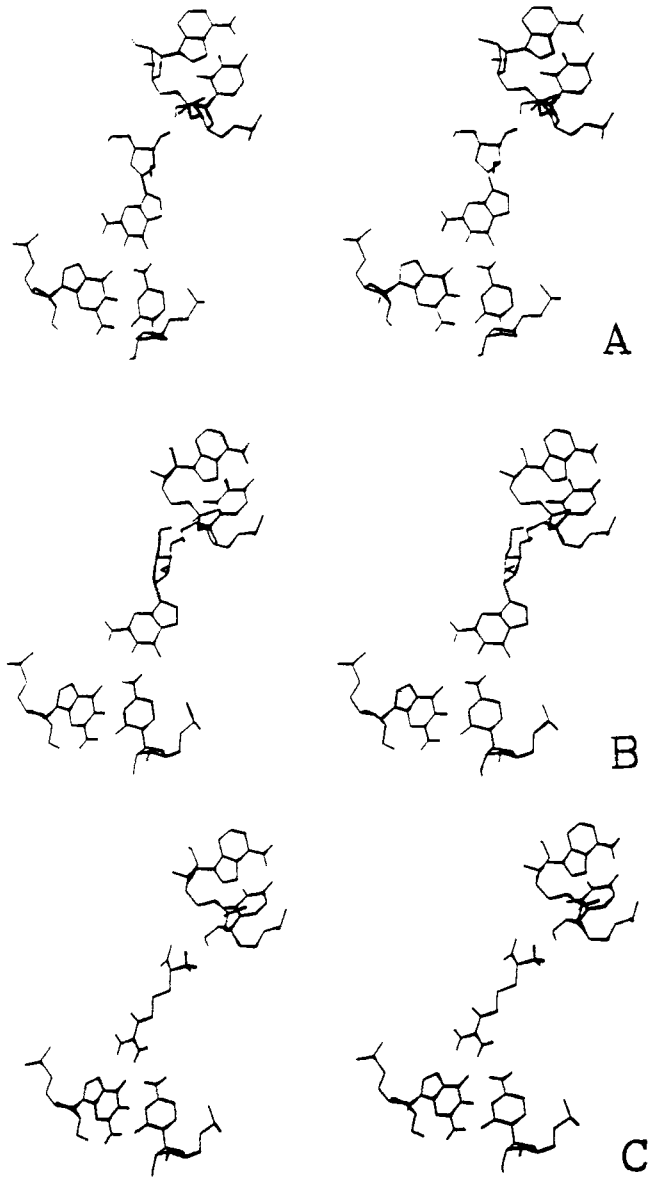

FIG. 4. Stereo drawings of the active site of the *nrdB* IVS with (A) the guanosine cofactor in its initial position, (B) the guanosine after MD simulation and minimization, and (C) with an arginine interacting with G_{458}-C_{526} after energy minimization.

need for further modifications of the model. To obtain a close contact (≈3Å) between the G-O3' and the U-1-O3' the cofactor was repositioned slightly in such a way that the hydrogen bond G-N2:G458-N7 is replaced by hydrogen bond G-O6:C526-N4 (FIG. 4A). During the MD simulation the guanosine moved even closer, by 1Å, to the cleavage site at the expense of losing its hydrogen bond to G458 (FIG. 4B). The guanosine also rotated slightly and moved a little bit out of the plane of G458 and C526, increasing its interactions somewhat with other nearby nucleotides. Nucleotides C457, A456 and U525 are the ones that are closest (apart from G458 and C526) to the guanosine base with A456 being close to the guanosine ribose as well. Of these U525 is conserved within group 1 and group 1A[4].

The final MD average structure of the IVS-Gua complex was used as a docking target (with the guanosine cofactor removed) for an arginine inhibitor[20]. The arginine replaced the guanosine in a similar way as proposed by Michel *et al.*[2], with hydrogen bonds between the two amino groups of the arginine sidechain and the G458 N7 and O6 atoms (FIG. 4C). This positions the arginine NH_3^+ close to the U-1 phosphate group, thus creating a stabilizing electrostatic interaction between the arginine and the RNA. This interaction is, however, not specific for arginine since it does not involve the sidechain.

It has been proposed[1] that the helix P1 is attached to the single stranded regions G516UGAU520 and U42AAU45 (nucleotides in bold face are conserved in group 1 and group 1A), which are indeed in direct contact with P1 in the present model: A44 is hydrogen bonded to U-1, U520 is hydrogen bonded to G-4 and to the sugar of C10.

Conclusions

A computer built model of a large RNA, consisting of three extended helices in the core of the self-splicing intron of the T4 *nrdB* gene, forms a stable structure which can be used to study interactions with various ligands. There is indeed a suitable cavity between helices P1 and P7 in which the cofactor guanosine, necessary for the splicing reaction, fits nicely, as does the inhibitor arginine. The geometry of this active site is such that the ligands form specific contacts with nucleotides in helix P7, in accordance with studies on the *Tetrahymena* intron by Michel *et al.*[2], while maintaining a close contact between the reacting groups of the intron and the guanosine.

Acknowledgements

The authors are grateful to the Department of Molecular Genetics, Karolinska Institutet, and to Dr. Ulf SKoglund, for the use of their CONVEX C210 computer. This work was supported by grants from the Swedish Natural Science Research Council, the Lisa and Johan Grönberg Foundation, the Knut and Alice Wallenberg Foundation, the O.E. and Edla Johansson Scientific Foundation, and the Magnus Bergvall Foundation.

REFERENCES

1 Kim S.H. and Cech T.R. (1987) Proc. Natl. Acad. Sci. USA **84**, 8788-8792.

2 Michel F., Hanna M., Green R., Bartel D.P. and Szostak J.W. (1989) Nature **342**, 391-395.

3 Burke J.M., Belfort M., Cech T.R., Davies R.W., Schweyen R.J., Shub D.A., Szostak J.W. and Tabak H.F. (1987) Nucl. Acids. Res. **15**, 7217-7221.

4 Cech T.R. (1988) Gene **73**, 259-271.

5 Sjöberg B.M., Hahne S., Mathews C.Z., Mathews C.K., Rand K.N. and Gait M.J. (1986) EMBO J. **5**, 2031-2306.

6 Gott J.M., Shub D.A. and Belfort M. (1986) Cell **10**, 81-87.

7 Chu F.K., Maley G.F., Maley F. and Belfort M. (1984) Proc. Natl. Acad. Sci. USA **81**, 3049-3053.

8 Shub D.A., Gott J.M., Xu M.Q., Lang B.F., Michel F., Tomaschewski J., Pedersen-Lane J. and Belfort M. (1988) Proc. Natl. Acad. Sci. USA **85**, 1151-1155.

9 Pedersen-Lane J. and Belfort M. (1987) Science **237**, 182-184.

10 Quirk S.M., Bell-Pedersen D., Tomaschewski J., Rueger W. and Belfort M. (1989) Nucl. Acids Res. **17**, 301-314.

11 Kim, S.H., Crystal Structure of Yeast tRNA Phe and General Structural Features of Other tRNAs. In TRANSFER RNA: Structure, Properties, and Recognition (Eds. Schimmel, P., Söll, D. & Abelson, J.), pp. 83-100, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1979.

12 Latham J.A. and Cech T.R. (1989) Science **245**, 276-282.

[13] Nilsson L., Åhgren-Stålhandske A., Sjögren A.S., Hahne S. and Sjöberg B.M. (1990) Biochemistry in press.

[14] Brooks B.R., Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S. and Karplus M. (1983) J. Comp. Chem. **4**, 187-217.

[15] Nilsson L. and Karplus M. (1986) J. Comp. Chem. **7**, 591-616.

[16] Singh U.C., Weiner S.J. and Kollman P. (1985) Proc. natl. Acad. Sci. USA **82**, 755-759.

[17] Shub D.A., Goodrich H.A., Gott J.M., Xu M.Q. and Scarlato V. (1988) J. Cell. Biochem. Suppl. **12D**, 30.

[18] Arnott, S., Campbell Smith, P.J. & Chandrasekaran, R., Handbook of Biochemistry and Molecular Biology (G.D. Fasman), 419, CRC Press, Cleveland, 1975.

[19] Inoue T., Sullivan F.X. and Cech T.R. (1986) J. Mol. Biol. **189**, 143-165.

[20] Yarus M. (1988) Science **240**, 1751-1758.